Claims 1, 4-10, 12-15, 18-21 and 24-26 were pending at the time of examination. Claims 15, 18-21 and 24-26 have been canceled. The Applicant is <u>not</u> conceding that the subject matter encompassed by claims 15, 18-21 and 24-26 prior to this amendment is not patentable over the art cited by the Examiner. Claims 15, 18-21 and 24-26 were cancelled in this amendment solely to facilitate expeditious prosecution of the remaining claims. The Applicant respectfully reserves the right to pursue additional claims, including the subject matter encompassed by claims 15, 18-21 and 24-26, as presented prior to this amendment, in one or more continuing applications. The Applicant respectfully requests reconsideration based on the foregoing amendments and these remarks.

## Claim Rejections – 35 U.S.C. § 112

Claims 1, 4-10, 12-15, 18-21 and 24-26 were rejected under 35 U.S.C § 112, first paragraph, as not reasonably providing enablement for non-iterative clustering. The Applicant respectfully disagrees for the following reasons.

It appears that the Examiner's and the Applicant's paragraph numbering is inconsistent, so for the sake of clarity, it is restated here that in the previous response, when the Applicant referred to paragraph [0015] for support for non-iterative processing, the Applicant was referring to the last paragraph on page 5 of the specification, which reads in its entirety:

> [0015] The present system is particularly advantageous in that it provides an efficient and computationally inexpensive way to analyze the characteristics of unknown data. <u>Furthermore, performance of the clustering method requires only two passes over the data.</u>

The Applicant is <u>not</u> referring to the paragraph numbered [0015] by the Examiner, which among other things states that "…reference numerals are reused..", as that paragraph clearly refers to the reference numerals used in the drawings in the instant application and not to any features of the claimed process. Paragraph [0015] in the specification as filed, and as reproduced above, clearly supports the non-iterative behavior recited in the claims. As was discussed in the previous office action response, this non-iterative behavior clearly distinguishes the Applicant's invention as claimed over conventional, iterative, K-means algorithms that require iterative recalculation of a group of "centroid" locations until the centroids' locations do not change between subsequent recalculations.

Furthermore, whereas the other sections referred to by the Examiner (paragraphs [0042]-[0046] in the Examiner's numbering, corresponding to paragraph [0031] in the specification, as filed), may suggest an iterative process, this is not what is recited in the claims. Claim 1, for example, merely recites a refinement step that describes a single pass only of refinement, i.e., merging the smallest cluster with a bigger one. Claim 1 does not include the "repeat" aspect that the Examiner underlined in Examiner's paragraph [0046] on page 4 of the most recent office action. It should be noted that such a single refinement step may be sufficient to reduce the number of clusters to a desired number of clusters.

It should be noted again that the refinement step in claim 1 describes refinement operations on already established clusters in order to reduce the number of established clusters (see, for example, the paragraph referred to by the Examiner as paragraph [0042]) and not cluster formation itself. Thus, these sections (numbered [0042]-[0046] by the Examiner) do provide adequate support for claim 1 and do not contradict the notion of non-iterative clustering recited in claim 1.

Consequently, the description fully enables the non-iterative clustering method of claim 1. Thus, it is respectfully requested that the rejection of claim 1 under 35 U.S.C § 112, first paragraph, be removed. For similar reasons, it is also requested that the rejection of dependent claims 4-10 and 12-14 be removed.

### Claim Rejections – 35 U.S.C. § 103

Claims 1, 4-10, 12-15, 18-21 and 24-26 were rejected under 35 U.S.C § 103(a) as being unpatentable over U.S. Patent No. 6,012,058 to Fayyad et al. (hereinafter "Fayyad") in view of U.S. Patent No. 6,636,862 to Lundahl et al. (hereinafter "Lundahl"). The Applicant respectfully traverses this rejection.

On a high level, Fayyad describes his techniques as "an intelligent sampling scheme that employs some theoretically justified criteria for deciding which data can be summarized and represented by a significantly compressed set of sufficient statistics, and which data items must be carried in computer memory, and hence occupying a valuable resource" (col. 2, lines 56-62). Fayyad further describes how "on any given iteration of the invention" the existing data samples are partitioned into a discard set (DS), a compression set (CS), and a retained set (RS), where only the retained set is kept in memory and the other two are only kept in the form of "representative sufficient statistics." That is, in contrast to the Applicant's invention as defined in claim 1, Fayyad does not even assign all records to clusters. In the interest of saving "valuable

resources", Fayyad only saves "representative sufficient statistics" pertaining to at least a portion of the records.

Fayyad's techniques use traditional K-means clustering (see, for example, col. 6, lines 6-63). The differences between the conventional, iterative, K-means clustering and the non-iterative techniques used in the Applicant's invention as claimed have been thoroughly discussed in previous office action responses. For example, in the K-means algorithm, the distance is iteratively determined from a centroid, which is, in a sense, a "moving target" as the position of the centroid changes in every iteration. In the Applicant's method, however, the deviation is determined with respect to the <u>characteristic value</u> (i.e., a mean or a median value), which does not change for a given attribute and a given set of records.

Furthermore, even if one were to entirely disregard the differences between the iterative nature of Fayyad and the non-iterative nature of the Applicant's invention, as defined in claim 1, it is clear that Fayyad does not anticipate or render obvious the Applicant's invention as defined in claim 1. For example, in addition to merely assigning "representative statistics" (and not records) to clusters as discussed above, Fayyad also fails to provide a key that "comprises an ordered list of the set of attributes and the deviations from the characteristic value of said each attribute". The sections cited by the Examiner to allegedly anticipate the "key" of claim 1 (i.e., col. 8, lines 31-45 and col. 17 of Fayyad) do not in fact provide a key. On the contrary, the cited sections describe "a second embodiment of the process of updating the dataset DS" (col. 8, lines 31-32), that is, the discard dataset, for which only "representative sufficient statistics" (i.e., no records) is kept.

Consequently, since no key exists in Fayyad, Fayyad cannot reasonably disclose or render obvious the claimed limitation of "combining the set of records based on the key into a clustering result that includes a plurality of clusters." The sections cited by the Examiner this time (i.e., col. 9, lines 45-56 and col. 23 of Fayyad) describe how to form a new compression dataset by "<u>combining subclusters</u> using hierarchical agglomerative clustering" based on a "spread criteria." This is an iterative merging process that operates on clusters that have already been formed. This is clearly different from doing an initial cluster assignment of records based on the key described in claim 1.

Even if one were to combine Fayyad and Lundahl, the combination would not remedy the deficiencies discussed above. Consequently, it is respectfully submitted that claim 1 is neither anticipated nor rendered obvious by the Fayyad and Lundahl combination. It is respectfully requested that the rejection of claim 1 under 35 U.S.C § 103(a) be withdrawn.

Claims 4-10 and 12-14 all depend directly or indirectly from claim 1. Thus, it is respectfully submitted that claims 4-10 and 12-14 are not anticipated nor rendered obvious for at least the reasons presented above with respect to claim 1, and it is respectfully requested that the rejection of claims 4-10 and 12-14 under 35 U.S.C § 103(a) be withdrawn.

## Conclusion

The Applicant believes that all pending claims are allowable and respectfully requests a Notice of Allowance for this application from the Examiner. Should the Examiner believe that a telephone conference would expedite the prosecution of this application, the undersigned can be reached at the telephone number set out below.

Respectfully submitted,
MOLLBORN PATENTS

/Fredrik Mollborn/

Fredrik Mollborn, Reg. No. 48,587

2840 Colby Drive
Boulder, CO 80305
(303) 459-4527